Semantic-based framework for personalised ambient media

Aroyo, L.M.; Bellekens, P.; Björkman, M.; Houben, G.-J.

published in
Multimedia Tools and Applications
2008

DOI (link to publisher)
10.1007/s11042-006-0078-3

citation for published version (APA)
Aroyo, L. M., Bellekens, P., Björkman, M., & Houben, G-J. (2008). Semantic-based framework for personalised ambient media. *Multimedia Tools and Applications*, *36*(1-2), 71-87. https://doi.org/10.1007/s11042-006-0078-3

Multimed Tools Appl (2008) 36:71–87
DOI 10.1007/s11042-006-0078-3

Semantic-based framework for personalised ambient media

Lora Aroyo · Pieter Bellekens · Martin Björkman · Geert-jan Houben

Published online: 21 December 2006

Abstract The paper proposes a semantic-based metadata framework for personalised interaction with TV media in a connected home context. Our approach allows the current home media centres to go beyond the simple concept of electronic programme guides and to offer the users a personalised media experience in an ambient home environment. The user's characteristics, preferences and context are used to personalise the user's experience of viewing and interacting with multimedia content on different heterogeneous devices. The TV-Anytime specification provides the basis for the metadata framework for handling content from IP, digital broadcast, and Blu-ray disc sources.

Keywords Digital broadcast · Blu-ray disc · Semantic-based metadata framework

1 Introduction

The information society is going digital to an even greater extent in the field of media. These changes bring new possibilities and challenges which affect the whole media chain: from content production, via distribution, to last but not least the end-user (the consumer). We describe in this paper our research focussing on the experience of the home user and the possibilities for connecting several digital media input channels and user devices into a connected media centre.

L. Aroyo (✉) · P. Bellekens · M. Björkman · G.-j. Houben
Department of Mathematics and Computer Science, Eindhoven University of Technology,
P. O. Box 513, 5600 MB, Eindhoven, The Netherlands
e-mail: l.m.aroyo@tue.nl

P. Bellekens
e-mail: p.a.e.bellekens@tue.nl

M. Björkman
e-mail: m.bjorkman@tue.nl

G.-j. Houben
e-mail: g.j.houben@tue.nl

As an effect of digitalising, new forms of home media are emerging as digital systems are converging. Different content, e.g., from TV, radio, music, homemade images and videos, are no longer bound to separate devices or to local storage, and the development of the Internet makes the media boundaries become less limiting. As envisioned by for instance IBM [2], the future media may become more pervasive and offer a more ubiquitous and immersive experience, as increasing technological sophistication brings new media environments. The transfer to digital content along with technologies and standards like DVB,[1] HDTV, voice over IP, Blu-ray[2] and TV-Anytime[3] create opportunities to bring new interactivity to the traditional TV concept and change it drastically. The television itself has not yet experienced a major revolution for the past fifty years, which constitutes a strong contrast to the Internet which has quickly evolved from mere textual information to multimedia content. We believe that using Semantic Web technology in the TV content interaction concept may provide changes from a traditional one-way communication to a two-way communication where the user changes from a passive viewer to a more active participant, and programme structures change from fixed to dynamic. Furthermore, the home users would, to a greater extent, also become content producers [2], thus breaking the traditional business model where companies and institutions are the sole content providers.

In this paper we try to identify requirements, opportunities and problems in home media centres and we propose an approach to address them by describing an intelligent home media environment. The major issues investigated are coping with the *information overflow* in the current provision of TV programs and channels and the *need for personalisation* for specific users by adapting to their age, interests, language abilities, and various context characteristics. The research presented in this paper is a collaboration between the Eindhoven University of Technology and the Philips Applied Technologies group. The work has been carried out within the ITEA-funded European project Passepartout, which also includes partners like Thomson, INRIA and ETRI.

In Section 2 we describe the motivation and problem in relation to related work, followed by an illustrative use case scenario in Section 3. The TV-Anytime packaging concept is described in Section 4 and serves as the background for understanding our proposed system architecture described in Section 5. The latter elaborates on an interoperable design and on semantic techniques for enabling intelligent context-aware personalisation. Section 6 describes the implementation followed by stating future work in Section 7.

2 Related work

We investigate the design of a home media architecture of connected devices that can provide access to a wide range of media sources, yet at the same time avoid an overflow of information for the user. In our demonstrator called *SenSee: Sensing the user for Personalized Access to TV content*, we aim to connect devices such as shared (large) screens, personal (small) handheld devices, hand gesture recognition and biosensor-based interfaces. This intends to go beyond the traditional limited solution of a single TV screen and simple remote control and thus creates the foundation for an ambient home environment to collect various data about the users and to subsequently use this data for

[1] http://www.dvb.org

[2] http://www.blu-ray.com

[3] http://www.tv-anytime.org

the personalization of his/her interaction with the TV content. Related work on connected homes can be found in the field of ambient intelligence, investigated for instance at the Philips HomeLab [7].

Regarding the *information overflow* aspect, we assume that the amount of available digital content will increase enormously with the current digital development [15]. Simple programme guides are thus likely to turn inefficient in terms of helping the user in choosing from an overwhelming amount of content [5, 17]. This creates a challenge for media systems to support the user by intelligent recommendations to find the most relevant and interesting programmes. Similar research focused on filtering for interactive TV systems in home environments have previously been done by e.g., Goren-Bar and Glinansky [8]. Here content filtering and user stereotypes were used for capturing and using user preferences.

Various researchers furthermore emphasize that there is a need for *personalisation* in dealing with a vast amount of TV content [1]. We believe that a personalization approach in home media centres is significant in order to handle the user's preferences as basis for the interaction both regarding content and devices. Since users differ in ages, interests, abilities and language preferences, it is important that these preferences can be represented in the system. For instance, an eight year old person will have very different favourite programs than an adult, and some user might want the movies to always be displayed on the biggest screen, but private content only on his or her handheld device. By creating a *user model* [11] for each user of the system, such personal preferences may be stored. This needs to capture both a *user profile*, with the user's preferences, and a *user context*, which describes the current situation that the user is in, for example whether the user is alone or with a group, what are the available devices at the moment, what is the time, what is the location etc. The models furthermore constitute a necessary requirement for enabling intelligent filtering of content to make *recommendations* [21]. By this we mean finding and suggesting content that should be interesting for the user, while filtering out unwanted or uninteresting information. Various filtering techniques for recommending movies have previously been explored by Masthoff [12], in which several user models are combined to create group filtering. Other related work is the PTVPlus online recommendation system for the television domain by O'Sullivan et al [17] and the Adaptive Assistant for Personalized TV by Yu and Zhou [22].

Apart from supplying semantic models of the user, it is also necessary to have sufficient metadata descriptions of the content. This constitutes the basis for content classification, i.e., sorting the content into different types like fiction, non-fiction, news, sports, etc. Intelligent search and filtering of content moreover benefit from metadata descriptions suitable for reasoning, to deduce new information and to enrich content search. Current ongoing research in this area by the W3C Multimedia Annotation on the Semantic Web Task Force has been described by Stamou et al [20].

Similar research as presented in this paper has furthermore been performed by Hong and Lim [10], who also propose using TV-Anytime for handling content in a personalised way. However, they focus on broadcasted content, whereas we also consider content from IP and removable media like Blu-ray. Furthermore, their solution for content search also uses keywords and user history to recommend content, although the architecture differs in that all processing occurs at a metadata server.

As will be described later, we propose the modelling of TV content with the use of ontologies. Relevant related work in this field can therefore be found in the domain of the Semantic Web. Necib and Freytag [16] have focused on using ontologies in query processing with a similar approach to ours, which aims at refining search queries with synonyms (and yet avoiding homonyms). However, we intend to take this one step further in our process as we also use other semantically related concepts and a measurement for semantic closeness.

3 Application scenario

In this section we describe a scenario to illustrate the target functionality of our demonstrator. The setting is in the home of a European, well-off family in the year 2010, which is living in a region outside their original parental background. While they wish for the children to integrate with the local community and live and learn from their neighbours, they also value their heritage (linguistic, cultural and religious), to effectively communicate with distant relatives and friends. The family consists of a mother, a father, a four year old child, a deaf nine year old, and a teenager. The parents are determined that the children should be effectively multi-lingual and multi-cultural, and will invest time to adapt the multimedia content in the home. They therefore act as media guides and to some extent teachers for their children, by selecting and adapting the content. Since the parents have immigrated to the region, they will have a different preference of content than the default local selection and they use their home media centre to include also programmes from their original home area, e.g., for news, music and movies. They may also choose to alter the language or subtitles of the content.

As the family gathers for a movie night together, the home media centre has suggested a movie that suits each family member's preference and interest. The mother has briefly scanned the story of the movie and discovered that the ending is in her opinion not suitable for the children. She therefore changes to an alternative ending. As they start the movie, they all together use a shared big screen. Although they use subtitles on this shared screen, this evening the deaf child also includes additional sign language on his personal small screen. The teenager on the other hand needs to practice her second language so her parents asked her this time to listen to an alternative language version with her headphones. Although they enjoy the movie together, the father also wants to follow a live soccer game broadcast, and therefore uses his own handheld screen to view this private video stream. The media devices in the home are all connected to the ambient home media environment.

4 TV-Anytime packaging

A content structure which goes beyond a fixed linear time structure and allows multiple languages, alternative versions, etc. puts high demands on the content model. It needs to have a dynamic structure, rich metadata, and be suitable for various media. We believe that the TV-Anytime phase 2 standard may serve as the basis in such requirements and we have built our demonstrator upon the TV-Anytime concepts. The standard focuses mainly on broadcast content, though we apply it here for content from IP and local disc sources. It organises content into packages; a package is an interconnected structure where each piece of content is referred to by a Content Reference Identifier, CRID, following an RFC standard [6]. This can be used for several purposes. To define locators, which give the actual location where the content is stored, to refer to the content's metadata, or for referring to some other set of CRIDs. The TV-Anytime package is thus a collection of related CRIDs. The data model of a package adopts the multi-level structure of the MPEG-21 Digital Item Declaration Language [4], i.e., a *container-item-component* structure, with some extensions.

For example, a language course structured as a package could be organised and divided into chapters and sections where each chapter or section is identified by a CRID. Figure 1 shows an example chapter consisting of sections of two exercises, where one has additional video clips. Each content element is not stored in the package itself, but is referenced using a locator. Thus some part can be distributed on a disc and an other via IP. The main content

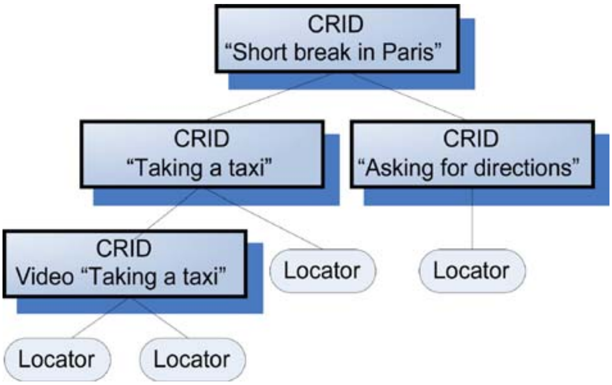

Fig. 1 Example CRID representation

of the language course could for instance be on a disc that the user has bought, extra interactive content and trailer for the next course may be distributed via IP. This packaging structure is very dynamic since parts can easily be modified or extended, for example the course could be extended with a new chapter by simply adding a CRID reference. Since packages are complex collections of CRIDs, they need to be resolved to discover which items are contained, as well as to get the locator(s) when viewing the actual content. This resolving process is performed by a *CRID Authority*. The response of such resolving request is an XML document containing a list of all CRIDs and locators in which it resolved.

The standard's metadata specification furthermore defines how content is described and classified—a fundamental feature for searching and filtering. The specification uses XML syntax for capturing different concepts. Apart from technical descriptions such as the screen aspect ratio and the number of audio channels there are possibilities of describing genre, synopsis, topic keywords, and language, etc. The genre description is a fine graded taxonomy structure, going from general concepts of fiction/non-fiction down to specific categories in the very leaf nodes. These are typically well known genres like comedy, drama, daily news, weather forecast, etc. that are used for content classification to sort programmes into categories. The TV-Anytime standard also describes the basic functions of how to access it via a *Metadata Service*. Our demonstrator uses these features and further maps additional semantics to the metadata in order to improve searching for and within packages.

5 Personalised home media centre

5.1 Design of personalised home media centre

We present in this section the architecture of the SenSee system and its components which can be seen in Fig. 2. We first begin with describing the services and the end-user environment before going into details of the central point of connection and control of TV-Anytime content.

Content Services represents in our design the different content distribution channels, where each has its own specific properties. Initially the optical disc will be the system's primary input channel of High Definition content. This is due to its high storage capacity, where we have chosen the Blu-ray disc as technology for this task. The IP channel can offer any type of content and has naturally the advantage of two-way communication, which makes IP well suited for interactive applications and distributing home made content to others. The SenSee

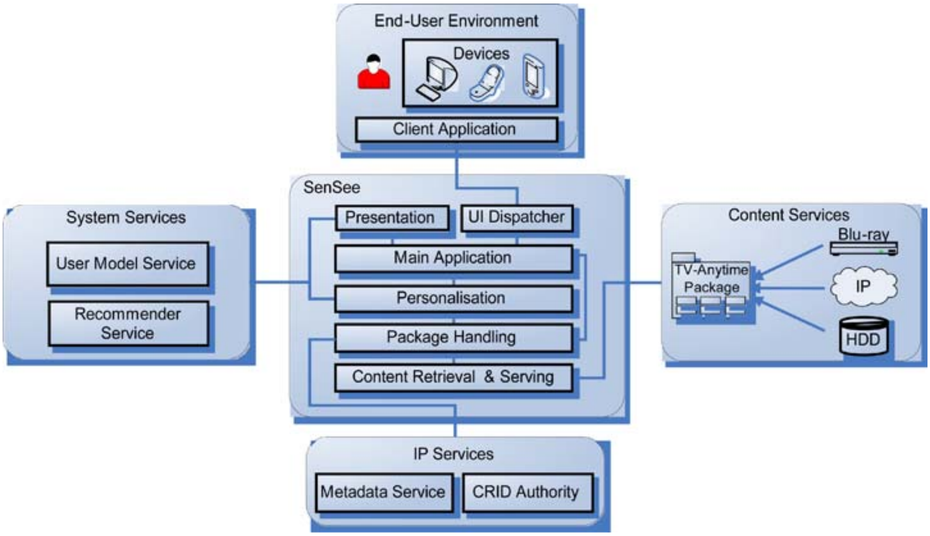

Fig. 2 Overview of central unit and external connections and services

demonstrator supports the most commonly used IP protocols of HTTP, FTP, streaming media and peer-to-peer technology. The hard disk drive (HDD) is used for local storage of content.

IP Services are external Web-based services that provide content metadata and descriptions of where content is located. This comprises the CRID Authority and the Metadata Service (see Section 4). However, in our system we have extended the Metadata Service to also handle more elaborate requests. The basic functionality to retrieve metadata for a CRID is useful when just the metadata, and not the whole package, needs to be retrieved: for example when browsing content, or showing some parts like title and synopsis to the user. However, our extension also enables searching for metadata which contains a set of keywords. This means searching for e.g., "French course" in all packages and retrieving a list of all which have metadata that contain these keywords. This significantly improves the search functionality.

System Services refer to the components that the system uses to handle the available content in an intelligent way. The *User Model Service* (UMS) is responsible for storing and updating information about the users, and the *Recommender Service* (RS) for using a set of filters to recommend content based on the user information. In other words, these are the main components to enable a personalised user experience. The user information can be accessed by sending queries to the UMS to retrieve e.g., a user's age, how a particular movie was ranked or, on a higher abstraction level, an estimation of the user's interest in a topic or genre like "sports". Such information is used by the Recommender Service to filter a large amount of information down to smaller result sets based on the user's interest perspective. Our content model, user model, and personalisation process involved in this are described in Sections 5.2 to 5.4.

The *End-User Environment* is the point where the user interacts with the system. Various devices, such as the TV screen, PC, handheld mobile or PDA are used. The central SenSee component identifies and keeps track of all connected devices and directs content streams simultaneously to multiple devices, according to the settings and user requests. Each user is furthermore identified and logged in to the system, in order to realise personalisation. Several users can be logged in and using the system in parallel, where each user may use one or multiple devices to interact with the system. The actual interface and functionality presented to the users are thus adapted for the devices used, with their own limitations and possibilities, and for the currently running applications.

SenSee connects to all services and user devices. In the middle (central) block in Fig. 2, starting at the bottom, the first two components provide the main functionality to handle content and content metadata. The *Content Retrieval & Serving* (CRS) component is responsible for the communication via the different connections of the Content Services. Retrieved content may be directly streamed through the system, cached or stored locally.

When downloading content, it communicates upwards in the component structure by events, to signal when the content has become available. As the name indicates, apart from retrieval, the Content Retrieval & Serving component also acts as a server for distributing local content, created by the users. Incoming external requests are handled and if accepted, the data transfer will be handled.

The *Package Handling* (PH) component communicates with the external IP Services concerning metadata descriptions and queries of package CRIDs. It enables the system to work with and parse packaged content as well as editing or making new content. Naturally, the Package Handling component keeps track of a CRID for locally stored or cached content to reduce communication. Furthermore, for any content that a user wants to publish, the Package Handling component communicates with the CRID Authority to receive a unique CRID and make this public. Thereafter, other systems can retrieve it via the Content Retrieval & Serving component.

Personalisation is described in detail in Section 5.4 and we therefore only mention here that it communicates with Package Handling, System Services and *Main Application* components. The latter is the session manager for keeping track of all active users, applications and devices. This means linking the applications to the content, as well as personalisation and presentation functions. The communication with client applications running on connected devices is based on a request-response model via a *User Interface Dispatcher*. In this process the interaction is directed via the personalisation, but there is also an option to skip this step as indicated in the figure, e.g., fin the case of the simple retrieval of a selected movie.

The last component, the *Presentation* component, provides functionality to influence how the interaction or the content should be presented for a particular set of devices, i.e., device specific adaptation. As it is linked to the System Services, the user model can be used in this process to retrieve a user's preferences regarding which device to use or regarding the look-and-feel using templates.

5.2 Semantically enriched content model

The content metadata previously described, is fundamental for searching and filtering of content. However, we imagine that due to the potentially vast amount of content, it is not enough to simply describe and classify the content, there must also be more intelligent ways of handling it. We therefore propose semantic knowledge models in addition to the TV-Anytime content model, which add possibilities for reasoning and deducing information about the content. The techniques we have used originate from the research area of the Semantic Web [3], where ontologies are used for modelling semantic relations between concepts. We have used OWL, the Web Ontology Language [13], to make a *semantically enriched content model* that serves two main purposes—providing us with the ability to:

- Reason and query the TV content
- Add semantic knowledge about the application domain to achieve intelligent behaviour of applications

To begin with, we have translated the TV-Anytime genre classification from XML into OWL in order to incorporate it into the system and enable querying. When translated into a *TV-Anytime content classification ontology* it is possible to use the structure and for example deduce that “archery” and “climbing” are both types of a “sports” genre. Without being able to use the linkage between the genre classification concepts, the applications will not “know” any semantic difference or connection between them. This facility is important for being able to group content into semantically related collections, which in turn is useful when presenting and navigating available content.

Furthermore, we have defined mappings from the TV-Anytime annotation elements to existing ontologies for time concepts, geography concepts, and lexical concepts which improve our possibilities to reason and query. Mapping time concepts to a time ontology [9], e.g., mapping the TV-Anytime annotation <PublishedTime> to corresponding time ontology concepts of year, month, day, timezone, etc. enables temporal reasoning over the data. By this we mean handling time intervals and the translation of expressions like “noon” or “evening”. Apart of having mappings from XML tags to ontology concepts, it is also useful when we are searching in text fields e.g., the synopsis. For example, a user might look for movies that take place in the 1970s. A simple search would only find those movies of which metadata explicitly mention some year in this decade (with a number 197*). However, by using concepts from the time ontology we can also find those movies that instead wrote “the seventies”. A geographical ontology, like the Teknowledge Ontology of Geography,[4] can likewise be used when searching for programs from “Europe”, where we extend the search with all the member countries to improve the results. When searching for programs about the region the user is located in, a geographical ontology gives us the area to use and possibly neighbouring cities. As a third ontology, a lexical ontology helps to find synonyms of terms. We have incorporated the WordNet [14] linguistic ontology in OWL.

Our approach is not limited to use only one ontology per domain. Geographical ontologies sometimes focus on listing countries while others specialize on defining orientations (like “westOf” or “isPartOf”). One time ontology may focus on time-zones while another concentrates on hour, day, month, year, etc. By combining the strengths of different ontologies, we can obtain a rich ontological structure. Furthermore, we intend to use specific *domain ontologies* that model typical topics, with one ontology for each major genre. For example, a “sports” ontology can model knowledge about sport equipment, famous players, and well-known competitions. Such domain ontologies are an additional source of knowledge that when searching for content can be used to semantically enrich the search. Our basic idea in this respect is to go beyond keyword matching, which usually is limited to finding only results which contain the exact keyword.

5.3 User modelling

For the purpose of personalisation we have defined a *User Model* (UM) to capture various concepts such as the user’s age, location, possible physical disabilities, interests, dislikings, etc. The model is divided into three parts: a model for the user profile, a second one for the user context and a third one for the user history.

The *User Profile* (UP) is used for capturing personal information of the user and his or her media preferences. The personal data comprises gender, age, home address, native language, other spoken languages and level of advancement, occupation, possible hearing or visual capacities, etc. Knowing this basic information enables content to be selected in accordance

[4] http://reliant.teknowledge.com/DAML/Geography.owl

with the correct age limit of the user, and finding content relevant for the user's local home area and in a language that the user understands etc. Preferred devices per type of content can also be stored. Over time the UP furthermore stores the user's content preferences, by linking a rating of liking/disliking to the content that was watched. Since content is described with the aid of TV-Anytime metadata, there is enough information to identify the programme or movie in order to be precise in remembering the programme when storing the ranking preference. Remembering how a user liked a particular movie like "Braveheart" for instance is valuable. Moreover it can be used for deriving conclusions about the user's preference for the genre of "war/drama" or for the main actor "Mel Gibson". As the content model has a searchable semantical model of the genres, it is possible to combine all ratings for some genre with the ratings of semantically related (sub) genres.

The *User Context* (UC) describes the present situation of the user. It comprises concepts of the time, the location, the devices, the audience and even the mood of the user. The purpose is to enable the system to achieve a more dynamic behaviour and adapt to the current situation. The time of the day may for instance be used in content filtering to affect the choice of device and content: for example one user may prefer a selection of news programs to be shown on the personal screen each morning. Furthermore, the currently available devices constitute a part of the user context. In the user context is also the fact whether the user is alone or part of a group audience that shares a device.

The *User History* (UH) contains information of the previous behaviour of the user. It stores every action that the user has taken and maintains a program list of the watched programs in combination with the time of watching. The purpose of this information is to support attempts to discover patterns in the user's behaviour. The program list is structured with a clear focus on time and comprises concepts of the week number, the day of the week, and exact start and end times. The model can thus be used to find which programs or types of programs are watched for example on a Tuesday evening. The User History can become rather extensive after a lot of user interaction with the system. The time structure will likely also contain a lot of useless, redundant and duplicate information after some time of use. Therefore it is necessary to have a process which can filter the structure periodically and keep only the valuable information. We have for this reason chosen to use two instances of the user history repository: a short term user history which records the information completely, and a more selective long term user history. The latter keeps all information which proved to be valuable after filtering the short term history, which happens at the end of a session when the user logs out. This process also updates the User Profile. The User Model thus develops and grows over time while the user interacts with the system. In an initial state it is however likely to suffer from a cold start problem, which is due to a low number of stored preferences, or from the fact that the first interactions influence the behaviour of the system too much. We attempt to avoid this phenomenon by employing simple stereotypical user templates which are selected at user registration by letting the user fill out a brief form with questions of user preferences.

5.4 Personalised content search

This section describes the personalised content search functionality of the SenSee Personalisation component. Our basic approach is to add personalisation in the step between the Main Application and Package Handling components when searching for content. This may occur when navigating through available content, when searching for something specific by entering keywords, or when asking the system to make a suggestion. In all cases, we aim at supporting the user by filtering the information based on the user's own perspective. The process affects the results found in the search in the following aspects:

- A smaller, more narrow result set is produced
- Results contain the items ranked as most interesting for the user
- Results contain the items most semantically related to any given keywords
- Searching goes beyond word matching search and also considers semantic related concepts
- Results are categorised with links to semantic concepts
- Semantic links can be used to show the path from search query to results

We illustrate this by stepwise going through the content search process as it is depicted in Fig. 3. Let us imagine the example that the user via the user application interface enters the keywords "*army 1940s*" and asks the system to search. This initial query expression of keywords ($k_1, ..., k_n$) is analysed in a query refinement process which aims at adding extra semantic knowledge. By using the set of available ontologies, we first search for modelled concepts with the same name as the keywords. We can in this case get hits in the history and time ontologies, where respectively "army" and "1940s" are found and thereby now known to belong to a history and time context. Second, since it is not sure that content metadata will use the exact same keywords, we add synonyms from the WordNet ontology, as well as semantically close concepts from the domain ontologies. In this case, apart from direct synonyms, a closely related concept such as "World War II" is found through a semantic link of "army" to "war" and "1940s" to "1945". Furthermore it links it to the geographical concept "West Europe" which in turn links to "Great Britain", "Germany" etc. However, this leads us to the requirement that the original keyword should be valued higher than related concepts. We solve this by adding a numerical value of semantic closeness, σ. In our initial algorithm, the original keywords and synonyms receive a σ value of 1.0, related ontology concepts within one node distance receive a value of 0.75 and those two nodes away a value of 0.5. Third, we enrich the search query by adding every occurrence we found together with a link to the corresponding ontology concept. The query is in that process refined to a new query expression of keywords ($k_1, ..., k_m$) ($m \geq n$), with links from keywords to ontology concepts ($\varsigma_1, ..., \varsigma_m$), and corresponding semantic closeness values ($\sigma_1, ..., \sigma_m$). Subsequently, the keywords in the query are mapped to TV-Anytime metadata items, in order to make a search request to the Metadata Service. From this content retrieval process the result is a collection of CRID references to packages which has matching metadata.

The next step in the process is *result filtering*, which aims at producing rankings of the search result in order to present them in an ordered list with the most interesting one at the top. Furthermore it performs the deletion of items in the list which are unsuitable, for example content with a minimum 18 years age limit for younger users. The deletion is a straightforward task of retrieving data on the user's parental guide limit or unwanted content types. The ranking consists of a number of techniques which estimate rankings from different perspectives:

- Keyword matching
- Content-based filtering
- Collaborative filtering
- Group filtering
- Context-based filtering

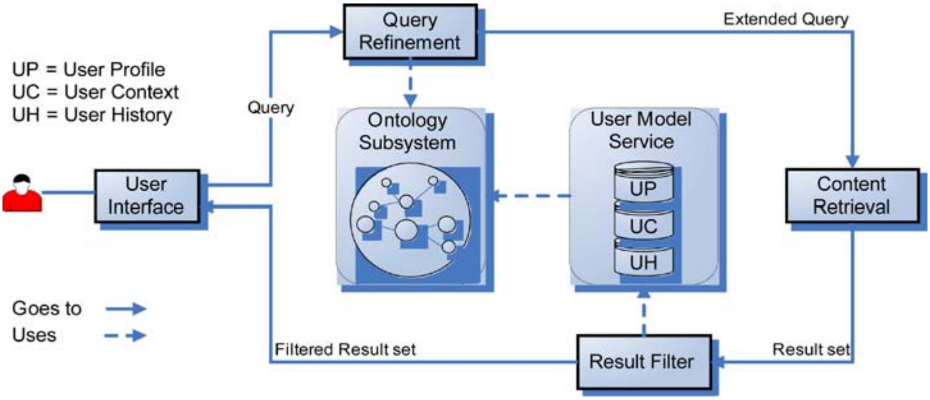

Fig. 3 Personalised package search process with semantic refinement

To begin with, packages are sorted based on a keyword matching value i.e., to what extent their metadata matched the keywords in the query. This can be calculated as average sum of matching keywords multiplied with the corresponding σ value, in order to adjust for semantic closeness. Content-based filtering [18] is furthermore used to predict the ranking of items that the User Model does not yet have any ranking of. This technique compares the metadata of a particular item and searches for similarities among the contents that already have a ranking, i.e., that the user has already seen. Collaborative filtering [19] is used to predict the ranking based on how other similar users ranked it. Furthermore, context-based filtering can be used to calculate predictions based on the user's context, as mentioned in Section 5.3. If there is a group of users interacting with the system together, the result needs to be adapted for them as a group. This can be done by for example combining the filtering of each individual persons to create a group filtering [11]. Finally, the ranking value from each technique is combined by summarising the products of each filter's ranking value and a filter weight.

5.5 Personalised presentations

In order to make the personalisation more transparent to the user, the path from original keyword(s) to resulting packages is shown when the results are presented. The synonyms and other semantically related terms are also made explicit to the user as feedback to the user that aims to avoid confusion when presenting the recommendation (e.g., when a movie with a different title is recommended than the original keyword given by the user). Since the links from keyword to related ontology concepts are kept we can present them in the user interface. Furthermore, we use them to group the result set, as well as in an earlier stage in the search process, when used to consult the user to find the appropriate context.

6 Implementation

The basic service-based architecture chosen for the system is illustrated in Fig. 4. It shows how the different SenSee services and content services connect.

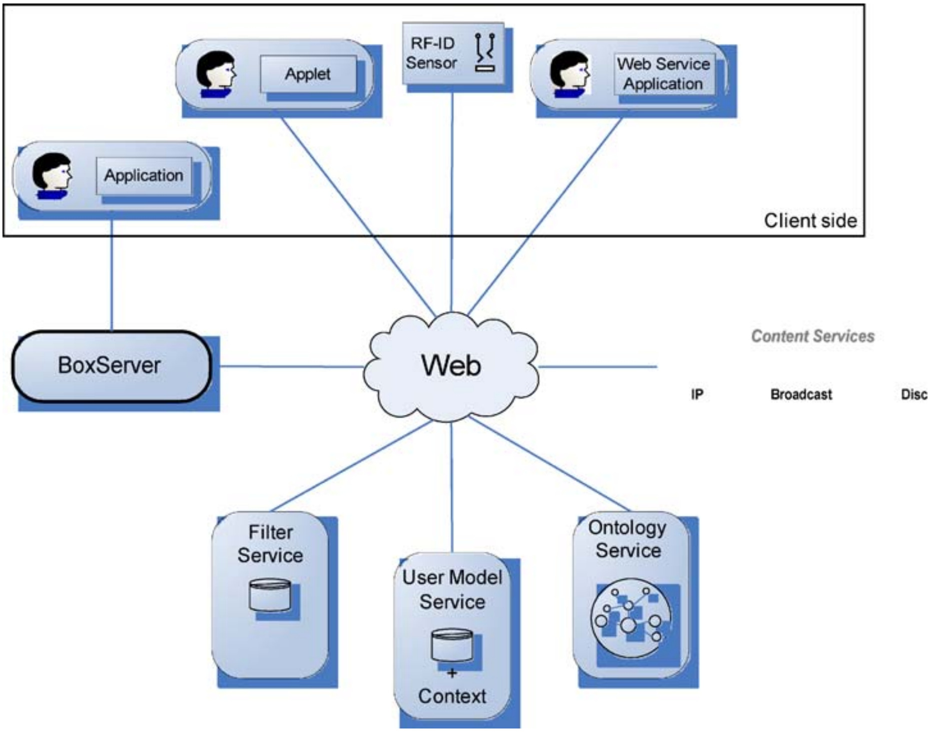

Fig. 4 SenSee client-server environment

A prototype of the system described has been developed and implemented in cooperation with Philips Applied Technologies. The fundamental parts of the IP and system services, content retrieval, packaging and personalisation are covered in this implementation. Our initial focus has been on realising the underlying TV-Anytime packaging concepts and personalisation, although not so much on the Blu-ray. Currently, geographical and synonym ontologies have been incorporated in the prototype. The integration of content ontologies is still under development. A simple end-user application for searching and viewing packages has also been made, which is exemplified in Fig. 5.

For test purposes a database with metadata on 500,000 movies from the Internet Movie Database (IMDB[5]) has been created and converted into TV-Anytime phase 2 packages. Our database also contains 1 million ratings of 4,000 movies from 3,000 users which have been imported from the EachMovie[6] dataset and are used by the prototype's first implementation of content recommendations. Our system currently allows a single user as well as multiple users to log in, where the system adapts to make recommendations for a single and a group of users correspondingly.

The User Model Service is an external service. This enables the use of multiple user models for collaborative filtering, improves the performance of the process, and allows for

[5] http://www.imdb.com

[6] http://research.compaq.com/SRC/eachmovie

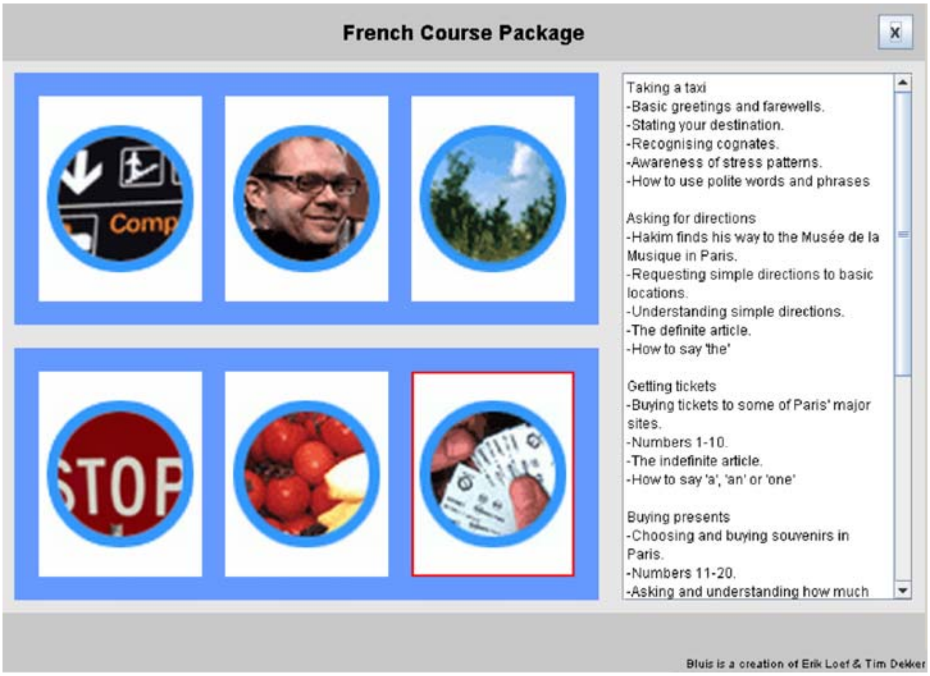

Fig. 5 Example view of package

doing an analysis of user behaviours in the large. Furthermore, the use of an external User Model Service gives possibilities for the user to access the User Model from other locations than in the own home. A positive effect is achieved for the so-called cold start problem, when the users use it on occasions, like visiting friends or on vacation, and get the same personalisation as at their own home.

However, it may be argued that it can lead to privacy or integrity problems if users are uncomfortable with the thought of having information about their behaviour stored somewhere outside their home system, no matter how encrypted or detached from the person's identity it can be done. These issues are currently outside the scope of the reported research.

The user devices that currently can be connected are a HDTV screen and a LIRC remote control which communicates through a JLIRC interface. Content Services can furthermore handle both local content as well as streaming content via IP. The implementation has mainly been made in Java, where connections of external services are realised by the Tomcat Web Server, Java Web Start, SOAP and RMI. The tools used for the application of semantic models are Sesame[7] and Protégé.[8]

The client application was originally developed as a stand-alone Java 5.0 application, including not only the client GUI interface but also administration views and pure test interfaces. Later the need of a Web-based client became clear to enable fast and easy access for external users. The SenSee Web Client, (see Fig. 6) was then implemented as an AJAX application. This is enabling us to provide a fluent Web experience without long waiting

[7] http://www.openrdf.org

[8] http://protege.stanford.edu

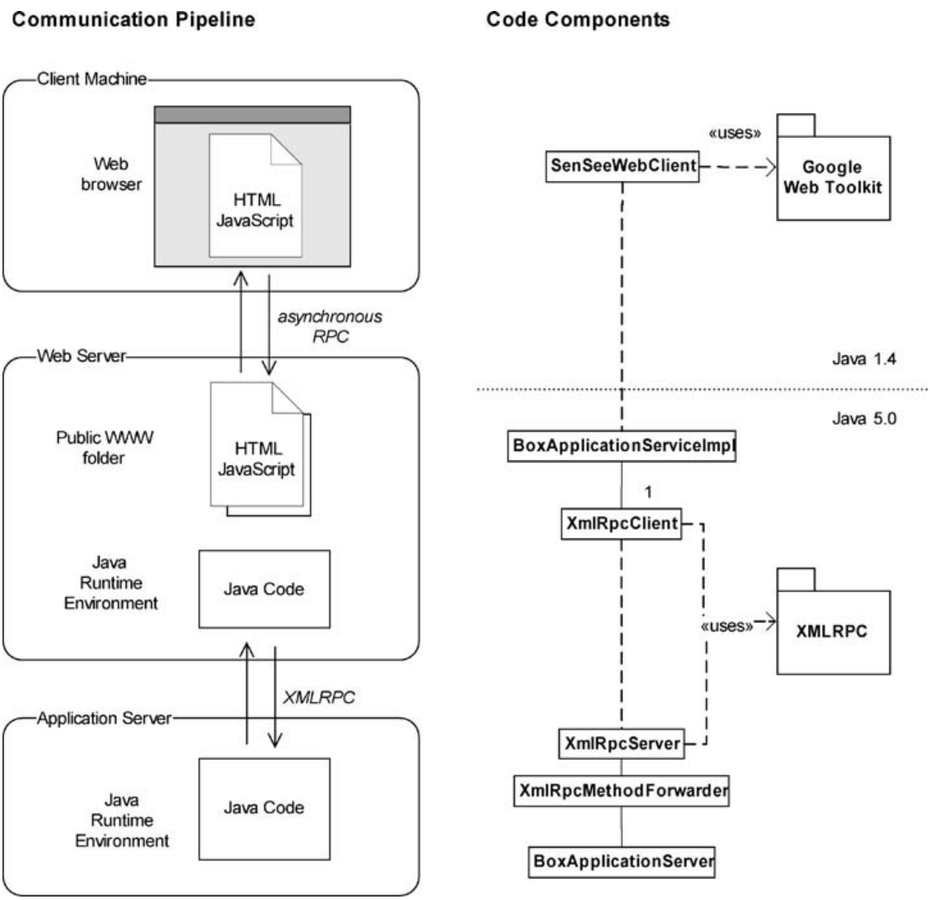

Fig. 6 SenSee web client architecture

times. The service is far from finished but the first navigation tests are running in connection with our main server (BoxServer). The pages themselves are built with the Google Web Toolkit.[9]

7 Conclusions

In this paper we described a scenario and an approach for a connected ambient home media management system that enables connections from both IP and Blu-ray, where users can view and interact via multiple rendering devices like TV screens, PDA, mobile telephone or other personal devices. The interaction, especially in content search, is supported by a semantics-aware and context-aware process which aims to provide a personalised user experience. This is important since users have different preferences and capabilities and the goal is to prevent an information overflow. We have presented a component architecture which covers content

[9] http://code.google.com/webtoolkit

retrieval, content metadata, user modelling, recommendations, and an end-user environment. Furthermore we have presented a semantically enriched content search process using TV-Anytime content classification and metadata. Our ultimate goal is to propose a fundamental platform that can be used further by applications and personalisation services. Our current work focuses on the exploration and design of appropriate user applications (e.g., personalized programme guide) which make use of the proposed architecture.

References

1. Ardissono L, Kobsa A, Maybury M (ed) (2004) Personalized digital television: targeting programs to individual viewers. Kluwer, Boston
2. Berman SJ (2004) Media and entertainment 2010. Open on the inside, open on the outside: The open media company of the future. Retrieved November 24, 2005, from http://www-03.ibm.com/industries/media/doc/content/bin/ME2010.pdf
3. Berners-Lee T, Hendler J, Lassila O (2001) The semantic web. Scientific American, New York
4. Bormans J, Hill K (2002) MPEG-21 Overview v.5. Retrieved November 24, 2005, from http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm
5. Chorianopoulos K (2004) What is wrong with the electronic program guide. Retrieved November 24, 2005, from http://uitv.info/articles/2004/04chorianopoulos
6. Earnshaw N, Aoki S, Ashley A, Kameyama W (2005) The TV-anytime Content Reference Identifier (CRID). Retrieved November 24, 2005, from http://www.rfc-archive.org/getrfc.php?rfc=4078
7. de Ruyter B, Aarts E (2004) Ambient intelligence: visualizing the future. In: Proceedings of the working conference on advanced visual interfaces. ACM, New York, pp 203–208
8. Goren-Bar D, Glinansky O (2004) FIT-recommending TV programs to family members. Comput Graph 28: 149–156 (Elsevier)
9. Hobbs J, Pan F (2004) An ontology of time for the semantic web. In: ACM Transactions on Asian Language Information Processing (TALIP), vol 3, Issue 1. ACM, New York
10. Hong B, Lim J (2005) Design and implementation of home media server using TV-anytime for personalized broadcasting service. In: Gervasi O, Gavrilova ML, Kumar V, Laganà A, Lee HP, Mun Y, Taniar D, Tan CJK (eds) Computational science and its applications—ICCSA 2005: conference proceedings, part IV, vol 3483. LNCS. Springer, Berlin Heidelberg New York, pp 138–147
11. Kobsa A (1990) User modeling in dialog systems: potentials and hazards. AI Soc 4(3): 214–231 (Springer-Verlag London Ltd)
12. Masthoff J (2004) Group modeling: selecting a sequence of television items to suit a group of viewers. User Model User-Adapt Interact 14: 37–85 (Kluwer)
13. McGuinnes DL, van Harmelen F (2004) OWL Web Ontology Language. Retrieved November 24, 2005, from http://www.w3.org/TR/owl-features
14. Miller GA (1995) WordNet: a lexical database for english. Commun ACM 38(11) (ACM)
15. Murugesan S, Deshpande Y (2001) Web engineering, software engineering and web application development. In: Murugesan S, Deshpande Y (eds) Web engineering, vol 2016. Lecture notes in computer science. Springer, Berlin Heidelberg New York
16. Necib CB, Freytag J-C (2005) Query processing using ontologies. In: Pastor O, Cunha JF e (eds) Advanced information systems engineering: 17th international conference, CAiSE 2005, vol 3520. Lecture notes in computer science. Springer, Berlin Heidelberg New York, pp 167–186
17. O' Sullivan D, Smith B, Wilson D, McDonald K, Smeaton A (2004) Improving the quality of personalized electronic program guide. User Model User-Adapt Interact 14: 5–36 (Kluwer)
18. Pazzani MJ (1999) A framework for collaborative, content-based and demographic filtering. Artif Intell Rev 13:393–408
19. Shardanand U, Maes P (1995) Social information filtering: algorithms for automating “Word of Mouth”. In: CHI ’95 Proceedings: conference on human factors in computing systems. pp 210–217
20. Stamou G, van Ossenbruggen J, Pan JZ, Schreiber G (2006) In: Smith JR (ed) Multimedia annotations on the semantic Web. IEEE Multimed 13(1):86–90
21. van Setten M (2005) Supporting people in finding information: hybrid recommender systems and goal-based structuring. Telematica Instituut fundamental research Series, No. 016 (TI/FRS/016). Universal, The Netherlands
22. Yu Z, Zhou X (2004) TV3P: an adaptive assistant for personalized TV. IEEE Trans Consum Electron 50 (1):393–399

Lora Aroyo is assistant professor of human-computer interaction at Free University Amsterdam, the Netherlands and part-time associated with Technische Universiteit Eindhoven, the Netherlands. Main research topics are (Semantic) Web technology, adaptation & personalization in Web applications, intelligent user interfaces, personalization for cultural heritage, interactive TV and e-learning. She (co-)organized, chaired, or served as PC member for numerous international events in the field, and serves as reviewer for several journals and conference related to these research topics.

Pieter Bellekens received his first M.Sc in Electronics at the De Nayer Institute in Belgium. Afterwards he moved to the Netherlands, more specifically Eindhoven, where he received an M.Sc in Mathematics and Computer Science in 2004. Staying at the university, he joined the group of Lora Aroyo to start his PhD program in the ITEA Passepartout project. Currently, he is involved in researching personalized search strategies for multimedia content via Semantic Web technologies and User Modeling for the ambient home environment.

Martin Björkman received his Master of Science degree in Computer Science in 2004 from the Royal Institute of Technology in Sweden. He is currently working as a Research Assistant at Eindhoven University of Technology where his research concerns access to digital multimedia using Semantic Web technology, user modeling, context adaptation and ambient intelligence. He is currently active in the ITEA project Passepartout which brings together academia and industry to create novel technology for the future digital TV.

Geert-Jan Houben is full professor of information systems at Vrije Universiteit Brussel in Belgium and part-time associated with Technische Universiteit Eindhoven in the Netherlands. He is heading the Hera research program on Web information systems, adaptation & personalization, and Semantic Web technologies. In Brussels he is codirector of the WISE research laboratory. Main research topics are (Semantic) Web technology, Web query and transformation languages, adaptation & personalization in Web applications, and Web engineering. He (co-)organized, chaired, or served as PC member for numerous events in the field, and serves as editorial board member or managing editor for journals in the area.